

## Устойчивое обеспечение сохранности в различных контекстах <sup>1</sup>

### 4. Коллективно создаваемый веб-контент

Социальные сети, краудсорсинг, блогосфера – эти виды контента возникли меньше десятилетия назад, но уже играют важную социальную, культурную и политическую роль в современной жизни. Все это типичные примеры коллективно создаваемого веб-контента – новой формы коммуникации и производства информации и знания не отдельным лицом или группой лиц, а коллективными усилиями (координированными или некоординированными) многих. Что из этого контента должно быть сохранено? На ком лежит за это ответственность? Очевидной является необходимость сохранить некоторую часть этого контента сегодня, оставить возможность отбора в будущем, предотвратить невосполнимую утрату.

Практически каждый аспект, который должен быть рассмотрен при принятии решения, обеспечивающего устойчивую сохранность, не имеет определенности. Кто выигрывает от обеспечения доступа к контенту в долгосрочной перспективе? В настоящее время мы знаем лишь, что представляет собой краткосрочный спрос. Кто владеет контентом? Многие вопросы о собственности и правах не имеют ответа, поскольку зачастую сайты являются результатом коллективного творчества множества людей с неопределенными правами на ту информацию, которую они размещают. Кто сохраняет контент? Сайты блог-хостинга, например, обычно не имеют определенной политики сохранения контента: мы не знаем, какой контент они сохраняют, а какой нет. И кто платит за это? Эти вопросы особенно беспокоят в том случае, когда сам сайт существует за счет усилий добровольцев, например, открытое программное обеспечение или веб-хостинг GeoCities <sup>2</sup>.

Ответы на эти вопросы не могут быть получены сразу. Устойчивая стратегия обеспечения сохранности коллективно создаваемого веб-контента, вероятно, будет отличаться от любой другой, выработанной для прочих цифровых материалов. Ответы на эти вопросы могут появиться с течением времени, в результате нескольких итераций, благодаря получению знания в процессе работы. Сегодня же важно правильно сформулировать вопросы, опробовать и смоделировать различные стратегии обеспечения сохранности. Некоторые действия следует предпринимать уже сегодня, например, вступать в партнерские отношения с создателями и владельцами

---

<sup>1</sup> Закінчення. Початок. див.: Шляхи розвитку української науки. – 2014. – № 3. – С. 70–74; Шляхи розвитку української науки. – 2014. – № 4. – С. 77–80; Шляхи розвитку української науки. – 2014. – № 5. – С. 102–108; Шляхи розвитку української науки. – 2014. – № 6. – С. 93–97; Шляхи розвитку української науки. – 2014. – № 7. – С. 92–95.

<sup>2</sup> См. <http://www.archiveteam.org/index.php?title=Geocities>

контента, что поможет пролить свет на природу ценности, стимулов, функции и ответственности тех, кто собирает и сохраняет эти материалы.

#### 4.1 Ценность и отбор

Несмотря на то, что прошло меньше десяти лет с момента появления этого контента, спрос на него очевиден: коллективно создаваемый веб-контент привлекает пользователей тем, что дает им возможность создавать и видоизменять информацию, взаимодействовать с ней, вносить свой вклад. Эта форма деятельности составляет основу всех жанров, которые мы проанализировали: блогосферы, социальных сетей (MySpace, Facebook, LinkedIn, Second Life) и общедоступного контента, создаваемого добровольцами (Slashdot, Wikipedia, Flickr, YouTube, SourceForge)<sup>3</sup>. Для этих жанров важен эффект накопления информации и сетевой эффект (взаимодействие с пользователями); чем больше добровольных авторов и пользователей, тем большую ценность для заинтересованных сторон приобретает ресурс, и тем больше обращений к сайту регистрируется.

##### Рассеянный или слабо выраженный спрос

Если оставить в стороне такие выдающиеся проекты, как Google Earth, Flickr, MySpace, Facebook, Wikipedia и некоторые активно читаемые блоги, особой ясности относительно того, какой именно контент следует сохранять, нет<sup>4</sup>. Отчасти это объясняется интерактивной и динамической природой подобных сайтов. Что значит «собирать блогосферу» – отбирать характерные блоги и представлять их? Что делать с блоггерами, которые не хотят, чтобы их материалы были сохранены, что бывает достаточно часто? Что значит сохранять Википедию (Wikipedia) – сохранять каждую статью и всю историю редактирования? Если и то, и другое, то возможно ли это сделать? Сегодняшняя практика архивирования и сохранения информации основывается на приблизительном консенсусе относительно границ информационного ресурса – где он начинается и где заканчивается, – так, чтобы можно было сохранить ресурс в окончательном виде или архивном формате. Но это не может стать моделью для контента, ценность которого основана на ссылках на различные сайты и отсылках к ним.

Несмотря на всю эту неопределенность, некоторые институты сегодня уже занимаются сохранением коллективно создаваемого веб-контента и лучше понимают, что надо делать, в процессе работы. Самым первым в 1996 г. собирать контент Всемирной Сети стал Архив Интернета (Internet

---

<sup>3</sup> См. <http://www.myspace.com>; <http://www.facebook.com>; <http://www.linkedin.com>; <http://secondlife.com>; <http://slashdot.org>; <http://www.wikipedia.org>; <http://www.flickr.com>; <http://www.youtube.com>; <http://sourceforge.net>.

<sup>4</sup> См. <http://earth.google.com>.

Archive). С этого момента многие библиотеки и архивы по всему земному шару собирают различные сегменты Интернета с целью сохранения информации. Очень часто они прибегают к стратегии отбора, собирая определенный разрез Сети большого объема и сохраняя его при сравнительно малом сопровождении вплоть до уточнения его долгосрочной ценности. Мы бы рекомендовали таким институтам возглавить деятельность по сохранению, убедив заинтересованные сообщества – создателей, пользователей, коллекционеров – в необходимости сформулировать краткосрочные цели отбора информации. Вслед за этим могут возникнуть партнерские отношения с хостингами, в рамках которых критерии отбора будут уточняться и будут определены технические требования к обеспечению сохранности.

**Рекомендация № 1:** Ведущим профильным организациям следует убедить заинтересованные стороны и экспертов в необходимости удовлетворения потребностей по отбору и сохранению коллективно создаваемого веб-контента.

#### 4.2 Стимулы к сохранению

##### Недостаточность стимулов

Мотивы, заставляющие добавлять контент на коллективные сайты, часто сводятся к быстрому получению выгоды. Производители ресурсов не всегда видят ценность долгосрочного сохранения. Фактически, ценность долгосрочного сохранения информации может быть очевидной для заинтересованной стороны, которая полностью изолирована от тех, кто создает контент, владеет им, вносит свой вклад в его изменение.

Помимо разбалансированности стимулов, очень сложно также определить, кто имеет право на сохранение контента. В сравнительно простом случае, когда речь идет о блоге, каждое сообщение («пост») может состоять из информации, связанной ссылками с материалами, которые защищены авторским правом. Блогеры вставляют в свои тексты материалы или отсылают к материалам, на использование которых у них нет определенных прав, и часто даже не пытаются получить на это разрешение. Государственный институт, имеющий стимул к сохранению таких блогов, безусловно, подумает не один раз, прежде чем примет решение по блогам со столь неопределенными правами. Эту ситуацию необходимо изменить. Одно из возможных решений – использование блогерами простых лицензий (таких, как Creative Commons), которые разъяснили бы их намерения в отношении архивирования контента третьей стороной – либо предоставление им неисключительных прав на сохранение контента на благо общества, либо запрет на архивирование третьей стороной.

Еще одно решение – стимулирование хост-сайтов к сохранению контента. Сайты, оказывающие услуги по хостингу, часто оказываются в конечном итоге единственной структурой, принимающей решения о долгосрочном размещении своего контента. Хостинги могут иметь сильные стимулы к сохранению материалов для частной выгоды, используя данные в целях получения прибыли. Однако и их можно мотивировать к сохранению контента для общественного блага. Разрешение на архивирование контента для некоммерческого использования, данное третьей стороне, не означает, что извлечение данных будет затруднено. Таким образом, можно рекомендовать хост-сайтам прибегнуть к практике лицензирования по умолчанию (без запроса), дающего третьей стороне неисключительные права на сохранение информации.

**Рекомендация № 2:** Создатели, авторы, хост-сайты могут облегчить задачу третьей стороне, если примут по умолчанию систему лицензирования, дающего третьей стороне неисключительные права на архивирование информации.

Второй ряд стимулов применим к деятельности некоторых крупных коммерческих хост-сайтов. Платформы доступа, на которых размещается и поддерживается коллективно создаваемый веб-контент – а это такие компании, как MySpace (принадлежащая Fox Interactive Media, филиалу News Corporation), а также YouTube (принадлежащая Google) – имеют сильные стимулы к сохранению контента на сайтах, с возможностью доступа к нему подписчикам и пользователям и извлечения интеллектуальных данных для рекламы и поддержки поиска, а также любых других целей по своему усмотрению. Напротив, маловероятно, что хостинговые компании имеют какой-либо стимул действовать в непонятных, размытых интересах будущих пользователей, например, социологов или специалистов по генеалогии. В этом случае, скорее всего, сильнейший стимул к сохранению информации найдется у третьей стороны, например, библиотек. Партнерские отношения между сайтом и архивирующей организацией могут быть выстроены таким образом, что пока материалы не потеряли своей коммерческой ценности, они используются и сохраняются сайтами, а материалы, представляющие долгосрочную ценность, по умолчанию передаются для сохранения.

Мы рекомендуем создавать при профильных организациях, а также сайтах, таких как Wikipedia, архивные филиалы, которые приняли бы на себя ответственность за сохранение этого контента на благо общества при финансовой поддержке. Оптимальным решением было бы создание консорциума организаций, распределивших между собой ответственность за сбор информации и достигающих эффективности за счет экономики

масштаба, а также принятия стандартов прозрачности и подотчетности в управлении этой информацией. Такой консорциум может быть создан под эгидой таких авторитетных организаций, как Библиотека Конгресса, Смитсоновский Институт, Королевская библиотека Нидерландов или Британская библиотека.

**Рекомендация № 3:** Регулирующие органы должны найти методы стимулирования, такие как субсидии, которые позволят хост-сайтам сохранять собственный контент или обращаться с целью его архивирования к третьей стороне на основе партнерства.

В общем и целом, создатели и пользователи этого контента, заинтересованные стороны мало озабочены обеспечением долгосрочного доступа к информации, которая, в сущности, является культурным наследием сегодняшнего дня, если они вообще осознают эту проблему. Это создает серьезную угрозу и диктует необходимость помощи со стороны государства. Государственные институты располагают наилучшими возможностями для долгосрочного сохранения цифровых материалов, обладающих высокой ценностью, которые никем не сохраняются или находятся под угрозой утраты по другой причине. В настоящее время в Соединенных Штатах Америки и в Европе этому препятствует один серьезнейший фактор – отсутствие четких полномочий на сбор в Интернете материалов для сохранения. Правительству Соединенных Штатов Америки мы рекомендуем предоставить такие полномочия до того, как Конгресс приступит к реформированию закона об авторском праве, и призываем к своевременной реформе законодательства.

**Рекомендация № 4:** Регулирующие органы должны принять незамедлительные меры по реформированию законодательства с тем, чтобы придать полномочия профильным организациям на сохранение веб-контента, находящегося под угрозой утраты.

#### 4.3 Функции, ответственность, финансирование

В отсутствие четкого распределения функций и ответственности за обеспечение сохранности коллективно создаваемого веб-контента, тем не менее, на низовом уровне предпринимаются усилия спасти контент, который может быть утрачен. Так, в 2009 г. группа добровольцев начала собирать контент хост-сайта Geocities (компания, которая разрабатывала и размещала в сети этот сайт, под таким же названием существовала с 1994 г., но внезапно прекратила свое существование). Тот же дух коллективных действий, который действует при создании подобных сайтов, поможет сохранить их, по крайней мере, в среднесрочной перспективе. Вопрос состоит в том, можно ли при помощи государственной политики или партнерских отношений с

институтами поддержать эти идущие снизу инициативы. Это предстоит исследовать лидерам процесса обеспечения сохранности веб-контента.

Для государственных институтов есть много возможностей, позволяющих им сыграть ключевую роль в сохранении коллективно создаваемого веб-контента. Так, Национальные библиотеки и другие ведущие профильные организации создали Международный консорциум по сохранности интернет-ресурсов (International Internet Preservation Consortium – ИПС), отчасти с целью определения функций и ответственности в сфере архивирования веб-материалов (распределение сфер ответственности происходит, главным образом, в соответствии с национальными интересами). Мы рекомендовали бы участникам ИПС и другим профильным организациям искать способы убеждения главных заинтересованных сторон, особенно владельцев и разработчиков сайтов коллективно создаваемого веб-контента, и развивать партнерские отношения с крупнейшими хост-сайтами.

Вопросов о том, что сохранять, как сохранять, кто будет за это платить, гораздо больше, чем ответов, поэтому важно убедить крупнейших поставщиков информации и хост-сайты в необходимости сохранять информацию, которую они обслуживают или которой они владеют. Также важно создать такие партнерские отношения, в рамках которых можно было бы исследовать технические, правовые и финансовые аспекты, обеспечивающие долгосрочный доступ к коллективно создаваемому контенту, представляющему ценность для будущих поколений.

**Рекомендация № 5:** Ведущие профильные организации должны развивать партнерские отношения с крупнейшими поставщиками информации с целью изучения технических, правовых и финансовых аспектов долгосрочного обеспечения сохранности контента.

#### Дефицит финансирования

Многие хост-сайты не имеют четкой модели финансирования, обеспечивающей им выживание на рынке, и в большинстве случаев борьба за выживание становится основной их задачей в ущерб стратегии долгосрочного сохранения контента. Коллективные усилия – обеспечение финансирования со стороны общества – добровольное или обязательное – будут необходимым условием выживания этих ресурсов в следующем десятилетии и далее, до тех пор, когда можно будет выявить ценность контента. В этих условиях возможности общественного финансирования включают в себя непосредственное субсидирование институтов, в задачи которых входит, в том числе, сохранение ресурсов, а также непосредственное финансирование программ сохранения частных ресурсов. Кроме того, в связи с тем, что некоторые участники процесса

получают государственную поддержку, уместным представляется наложение соответствующих обязанностей – выполнение которых и гарантирует государственную поддержку. В этом нет ничего нового: именно таким образом музеи и библиотеки находят средства на сохранение традиционных культурных материалов.

#### 4.4 Будущее

В большинстве случаев спрос (и, следовательно, предложение) определен лишь для сегодняшнего доступа. Сформулировать убедительное ценностное предложение в отношении долгосрочной сохранности только предстоит. Когда появится такое ценностное предложение, вполне возможно, что оно будет достаточно рассеянным и распределенным между создателями, авторами и пользователями ресурса. В связи с этим, сохранение его, вероятно, потребует коллективных усилий вплоть до того момента, как возникнет убедительное ценностное предложение и достаточные стимулы, побуждающие участников процесса действовать своевременно, до того момента, как все заинтересованные в сохранении стороны определятся со своими функциями и сферой своей ответственности. В настоящее время мы находимся лишь на начальном этапе разработки стратегии сохранения коллективно создаваемого веб-контента, когда эта задача стоит перед каждым участником, но ни один из них не готов ее решать. Тем не менее, мы видим, что на низовом уровне предпринимаются первые усилия, ведущие профильные институты примеряют на себя различные роли, используют разные модели финансирования, предпринимают меры, оставляющие людям возможность в будущем осознать долгосрочную ценность коллективно создаваемого веб-контента.

<b>БЛОК 4. Программа действий в сфере коллективно создаваемого веб-контента</b>
<ul style="list-style-type: none"><li>• <b>Рекомендация № 1:</b> Ведущим профильным организациям следует убедить заинтересованные стороны и экспертов в необходимости изучения удовлетворения потребностей по отбору и сохранению коллективно создаваемого веб-контента.</li><li>• <b>Рекомендация № 2:</b> Создатели, авторы, хост-сайты могут облегчить задачу третьей стороне, если примут по умолчанию систему лицензирования, дающего третьей стороне неисключительные права на архивирование информации.</li><li>• <b>Рекомендация № 3:</b> Регулирующие органы должны найти методы стимулирования, такие как субсидии, которые позволят хост-сайтам сохранять собственный контент или обращаться с целью его архивирования к третьей стороне на основе партнерства.</li><li>• <b>Рекомендация № 4:</b> Регулирующие органы должны принять незамедлительные меры по реформированию законодательства с тем, чтобы придать полномочия профильным организациям на сохранение веб-контента, находящегося под угрозой утраты.</li><li>• <b>Рекомендация № 5:</b> Ведущие профильные организации должны развивать партнерские отношения с крупнейшими поставщиками информации с целью изучения технических, правовых и финансовых аспектов долгосрочного обеспечения сохранности контента.</li></ul>

#### 4.5 Выводы

Это исследование современного ландшафта в сфере обеспечения сохранности цифровой информации в контексте четырех видов контента показало, насколько разнообразным может быть использование цифровых материалов, как разнообразна мотивация участников процесса и заинтересованных сторон. Тем не менее, независимо от самого контента, принятие решений имеет общие для всех ограничения, что объясняется тем, что все цифровые материалы имеют общие базовые свойства. В зависимости от контекста принимаются компромиссные решения. Однако во всех случаях краткосрочная программа действий должна предусматривать обеспечение доступа к информации в будущем... *(Устойчивая экономика для цифровой планеты: обеспечение долговременного доступа к цифровой информации. Итоговый отчет Рабочей группы по устойчивому обеспечению долговременной сохранности и доступа к цифровой информации [перевод с англ.]. – М.: МЦБС, 2013. – С. 142–153).*